ENHANCED ACCUMULATION OF TREHALOSE IN PLANTS

FIELD OF THE INVENTION

The invention relates to a method for the production of trehalose in plant cells, and plants. The invention is particularly related to a method for increasing the levels of trehalose accumulation in plants capable of producing trehalose. The invention further comprises higher plants, preferably *Angiospermae*, and parts thereof, which as a result of such methods, contain relatively high levels of trehalose. The invention further relates to plant cells, plants or parts thereof according to the invention obtained after processing thereof.

STATE OF THE ART

Trehalose is a general name given to D-glucosyl D-glucosides which comprise disaccharides based on two α-,α,ß- and ß,ß-linked glucose molecules. Trehalose, and especially α-trehalose 1-(O-a-D-glucopyranosyl)-1'-O-α-D-glucopyranose) is a widespread naturally occurring disaccharide. However, trehalose is not generally found in plants, apart from a few exceptions, such as the plant species *Selaginella lepidophylla (Lycophyta)* and *Myrothamnus flabellifolia*. Apart from these species, trehalose is found in root nodules of the *Leguminosae (Spermatophytae, Angiospermae)*, wherein it is synthesized by bacteroids; the trehalose so produced is capable of diffusing into the root cells. Apart from these accidental occurrences, plant species belonging to the *Spermatophyta* apparently lack the ability to produce and/or accumulate trehalose.

In International patent application WO 95/01446, filed on June 30, 1994 in the name of MOGEN International NV, a method is described for providing plants not naturally capable of producing trehalose with the capacity to do so. The method comprises introducing into the cells of said plants a recombinant polynucleotide encoding trehalose phosphate synthase under the control of regulatory elements necessary for expression of said recombinant DNA in plant cells. In one embodiment, tobacco and potato plants had been transformed with a recombinant polynucleotide encoding TPS from *E. coli*, under the control of the CaMV 35S RNA promoter. Levels of trehalose accumulation in these plants tended

to be rather low.

In spite of the absence of trehalose as a substrate in most higher plant species, the occurrence of trehalose-degrading activity has been reported for a considerable number of higher plant species, including those known to lack trehalose. The responsible activity could be attributed to a trehalase enzyme.

Reports suggest that trehalose, when fed to plant shoots grown *in vitro* is toxic or inhibitory to the growth of plant cells (Veluthambi K. *et al.*, 1981, Plant Physiol. 68, 1369-1374). Plant cells producing low trehalase levels were found to be generally more sensitive to the adverse effects of trehalose, than plants exhibiting a higher level of trehalase activity. Trehalose-analogs, such as trehalose-amines were used to inhibit trehalase activity in shoots, making it possible to study the effects of trehalose fed to plant cells. Plant shoots which produce relatively high amounts of trehalase were adversely affected by the addition of trehalase inhibitors. Inhibition of trehalase activity in homogenates of callus and suspension culture of various *Angiospermae* using Validamycin is disclosed by Kendall *et al.*, 1990, Phytochemistry 29, 2525-2582.

It is an object of the present invention to provide plants and plant parts capable of producing and accumulating trehalose, while keeping any adverse effects that may arise from the accumulation of trehalose within acceptable limits.

SUMMARY OF THE INVENTION

The invention provides a process for producing trehalose in plant cells capable of producing trehalase by growing plant cells having the genetic information required for the production of trehalose and trehalase, or cultivating a plant or a part thereof comprising such plant cells, characterised in that said plant cells are grown, or said plant or a part thereof, is cultivated in the presence of a trehalase inhibitor. Preferred plants or plant parts or plant cells have been genetically altered so as to contain a chimeric trehalose phosphate synthase gene in a plant expressible form. According to one embodiment said trehalose phosphate synthase gene comprises an open reading frame encoding trehalose phosphate synthase from *E. coli* in plant expressible form.

According to a further aspect of the invention, plants have been genetically altered so as to produce trehalose preferentially in certain tissues or parts, such as (micro-)tubers of potato. According to one embodiment the open reading frame encoding trehalose phosphate synthase from *E. coli* is downstream of the potato patatin promoter, to provide for preferential expression of the gene in tubers and micro-tubers of *Solanum tuberosum*.

According to another aspect of the invention the plants are cultivated *in vitro*, for example in hydroculture.

According to another preferred embodiment said trehalase inhibitor comprises validamycin A in a form suitable for uptake by said plant cells, preferably in a concentration between 100 nM and 10 mM, preferably between 0.1 and 1 mM, in aqueous solution.

Equally suitable said trehalase inhibition can be formed by transformation of said plant with the antisense gene to the gene encoding the information for trehalase.

Also suitable as trehalase inhibitor is the 86 kD protein from the american cockroach (*Periplaneta americana*). This protein can be administered to a plant in a form suitable for uptake, and also it is possible that the plants are transformed with DNA coding for said protein.

The invention further provides plants and plant parts which accumulate trehalose in an amount above 0.01 % (fresh weight), preferably of a *Solanaceae* species, in particular *Solanum tuberosum* or *Nicotiana tabacum*, in particular a micro-tuber of *Solanum tuberosum* containing trehalose.

The invention also comprises the use of a plant, or plant part, according to the invention for extracting trehalose, as well as the use thereof in a process of forced extraction of water from said plant or plant part. According to yet another embodiment of the invention a chimaeric plant expressible gene is provided, comprising in sequence a transcription initiation region obtainable from a gene, preferentially expressed in a plant part, particularly the patatin gene from *Solanum tuberosum*, a 5'-untranslated leader, an open reading frame encoding a trehalose phosphate synthase activity, and downstream of said open reading frame a transcriptional terminator region.

According to yet another embodiment of the invention a chimaeric plant expressible gene is provided, comprising in sequence a transcription initiation region obtainable from a gene, preferentially expressed in a plant part, particularly the patatin gene from *Solanum tuberosum*, a 5'-untranslated leader, an open reading frame encoding a trehalase coupled in the antisense orientation, and downstream of said open reading frame a transcriptional terminator region. A preferred plant expressible gene according to the invention is one wherein said transcriptional terminator region is obtainable from the proteinase inhibitor-II gene of *Solanum tuberosum*. The invention also provided vectors and recombinant plant genomes comprising a chimaeric plant expressible gene according to the invention, as well as a plant cell having a recombinant genome, a plant or a part thereof, consisting essentially of cells. A further preferred plant species according to this aspect is *Solanum tuberosum*, and a micro-tuber thereof.

The invention further provides a process for obtaining trehalose, comprising the steps of growing plant cells according to the invention or cultivating a plant according to the invention and extracting trehalose from said plant cells, plants or parts.

The following figures further illustrate the invention.

DESCRIPTION OF THE FIGURES

Figure 1. Schematic representation of binary vector pMOG799.

Figure 2. Schematic representation of binary vector pMOG845.

Figure 3. Schematic representation of parts of the sucrose and starch biosynthetic pathways in plant sink tissues. The figure shows that carbohydrate produced in the leaf by photosynthesis is transported via the phloem tissue in the form of sucrose. Upon entering the sink it is unloaded by a membrane bound invertase activity to yield the monosugars glucose and fructose. By the action of a number of enzymatic steps these monosugars are converted to starch and/or sucrose as roughly shown here. The glucose metabolites G6P and UDPG are believed to be used as the substrates for the TPS-enzyme engineered into the plant by introduction of the plant expressible otsA gene. The figure shows how the amount of UDPG and G6P available as substrate is increased by reducing the levels of the enzymes SPS and AGPase. Their inhibition is marked with a cross.

Figure 4. Alignments for maximal amino acid similarities of neutral trehalase from *S. cerevisiae* with periplasmatic trehalase from *E. coli*, small intestinal trehalas from rabbit and trehalase from pupal midgut of the silkworm, *Bombyx mori*. Identical residues among all trehalase enzymes are indicated in ***bold italics*** typeface. Conserved regions of the amino acid sequences were aligned to give the best fit. Gap's in the amino acid sequence are represented by dashes.
Positions of degenerated primers based on conserved amino acids are indicated by dashed arrows.

DETAILED DESCRIPTION OF THE INVENTION

According to the present invention it has been found that the accumulation of an increased level of trehalose in plants and plant parts is feasible, without causing too drastic effects on the viability of the plant or plant parts. This important finding can be exploited by adapting plant systems to produce and/or accumulate high levels of trehalose at lower cost.

According to one embodiment of the invention the accumulation of increased levels of trehalose is achieved by inhibiting endogenous trehalases. Inhibition of trehalases can be performed basically in two ways: by administration of trehalase inhibitors exogenously, and by the production of trehalase inhibitors endogenously, for instance by transforming the plants with DNA sequences coding for trehalase inhibitors.

This inhibition can be equally well applied to plants which are transformed with enzymes which enable the production of trehalose, but also to plants which are able to synthesize trehalose naturally.

According to this first embodiment of the invention, trehalase inhibitors are administered to the plant system exogenously. Examples of trehalase inhibitors that may be used in such a process according to the invention are trehazolin produced in *Micromonospora*, strain SANK 62390 (Ando *et al.*, 1991, J. Antibiot. 44, 1165-1168), validoxylamine A, B, G, D-*gluco*-Dihydrovalidoxylamine A, L-*ido*-Dihydrovalidoxylamin A, Deoxynojirimycin (Kameda *et al.*, 1987, J. Antibiot. 40(4), 563-565), 5-*epi*-trehazolin (Trehalostatin) (Kobayashi Y. *et al.*, 1994, J. Antiobiot. 47, 932-938), castanospermin (Salleh H.M. & Honek J.F. March 1990, FEBS 262(2), 359-362) and the 86kD protein from the american cockroach (*Periplaneta*

americana) (Hayakawa *et al.*, 1989, J. Biol. Chem. 264(27), 16165-16169). A pref rred trehalase inhibitor according to the invention is validamycin A (1,5,6-trideoxy-3-o-ß-D-glucopyranosyl-5-(hydroxymethyl)-1-[[4,5,6-trihydroxy-3-(hydroxymenthyl)-2-cyclohexen-1-yl]amino]-D-chiro-inositol). Trehalase inhibitors are administered to plants or plant parts, or plant cell cultures, in a form suitable for uptake by the plants, plant parts or cultures. Typically the trehalase inhibitor is in the form of an aqueous solution of between 100 nM and 10 mM of active ingredient, preferably between 0.1 and 1 mM. Aqueous solutions may be applied to plants or plant parts by spraying on leaves, watering, adding it to the medium of a hydroculture, and the like. Another suitable formulation of validamycin is solacol, a commercially available agricultural formulation (Takeda Chem. Indust., Tokyo).

Alternatively, or in addition to using exogenously administered trehalase inhibitors, trehalase inhibitors may be provided by introducing the genetic information coding therefor. One form of such in-built trehalase inhibitor may consist of a genetic construct causing the production of RNA that is sufficiently complementary to endogenous RNA encoding for trehalase to interact with said endogenous transcript, thereby inhibiting the expression of said transcript. This so-called "antisense approach" is well known in the art (*vide* inter alia EP 0 240 208 A and the Examples to inhibit SPS disclosed in WO 95/01446).

A gene coding for trehalase has been isolated from a potato cDNA library and sequenced. The predicted amino acid sequence of trehalose as shown in SEQIDNO:10 is derived from the nucleotide sequence depicted in SEQIDNO: 9. As is well known in the biological arts amino acid sequences of equivalent enzymes can differ between species. It is emphasized that the difference between the known trehalase sequences and plant trehalase sequence makes it very questionable if such trehalase sequence used in an antisense approach is capable of inhibiting trehalase expression *in planta*.

Of course the most preferred embodiment of the invention is obtained by transforming a plant with the antisense trehalase gene which matches exactly with the endogenous trehalase gene. However, sequences which have a high degree of homology can also be used. Thus, the antisense trehalase gene to be used for the transformation of potato will be directed against the nucleotide sequence depicted in SEQIDNO: 9.

It is usually enough to express only part of the homologous gene in the antisense orientation, in order to achieve effective inhibition of expression of the endogenous trehalase (*vide* Van der Krol *et al.*, 1990, Plant Molecular Biology, 14, 457-466).

Trehalase gene sequences of other plants can be elucidated in two different ways. One of the strategies is to use the isolated potato cDNA clone as a probe to screen a cDNA library containing the cDNA of the desired plant species. Positive reacting clones can then be isolated and subcloned into suitable vectors.

A second strategy to identify such genes is by purifying the proteins which are involved in trehalose degradation.
An example for such a strategy is the purification of a protein with acid invertase activity from potato (Solanum tuberosum L.) tubers (Burch *et al.*, Phytochemistry, Vol.31, No.6, pp. 1901-1904, 1992). The obtained protein preparation also exhibits trehalose hydrolysing activity. Disaccharide hydrolysing activity of protein preparations obtained after purification steps can be monitored as described by Dahlqvist (Analytical Biochemistry 7, 18-25, 1964).

After purifying the protein(s) with trehalose hydrolysing activity to homogeneity, the N-terminal amino acid sequence or the sequence of internal fragments after protein digestion is determined. These sequences enable the design of oligonucleotide probes which are used in a polymerase chain reaction (PCR) or hybridization experiments to isolate the corresponding mRNAs using standard molecular cloning techniques.

An isolated cDNA encoding a trehalose degrading enzyme is subsequently fused to a promoter sequence in such a way that transcription results in the synthesis of antisense mRNA.

Another form of such an in-built trehalase inhibitor may consist of a genetic construct causing the production of a protein that is able to inhibit trehalase activity in plants. A proteinaceous inhibitor of trehalase has been isolated and purified from the serum of resting adult american cockroaches (*Periplaneta americana*) (Hayakawa et al., *supra*). This protein, of which the sequence partly has been described in said publication, can be made expressable by isolation of the gene coding for the protein, fusion of the gene to a suitable promoter, and transformation of said fused gene into the plant according to standard molecular

biological methods.

A promoter may be selected from any gene capable of driving transcription in plant cells.

If trehalose accumulation is only desired in certain plant parts, such as potato (mini-)tubers, the trehalase inhibitory DNA construct (*e.g.* the antisense construct) comprises a promoter fragment that is preferentially expressed in (mini-)tubers, allowing endogenous trehalase levels in the remainder of the plant's cells to be substantially unaffected. Thus, any negative effects of trehalose to neighbouring plant cells due to trehalose diffusion, is counteracted by unaffected endogenous trehalase activity in the remainder of the plant.

In the Example illustrating the invention, wherein trehalose phosphate synthase is produced under the control of the patatin promoter fragment, also the trehalase-inhibitory construct may comprise a promoter fragment of the patatin gene.

Mutatis mutandis if trehalose is to be accumulated in tomato fruit, both a plant expressible trehalose phosphate synthase gene, which is at least expressed in the tomato fruit is to be used, as well as a plant expressible trehalase-inhibitory DNA construct, which should be expressed preferentially in the fruit, and preferably not, or not substantially, outside the fruit. An example of a promoter fragment that may be used to drive expression of DNA-constructs preferentially in tomato fruit is disclosed in EP 0 409 629 A1. Numerous modifications of this aspect of the invention, that do not depart from the scope of this invention, are readily envisaged by persons having ordinary skill in the art to which this invention pertains.

An alternative method to block the synthesis of undesired enzymatic activity such as caused by endogenous trehalase is the introduction into the genome of the plant host of an additional copy of said endogenous trehalase gene. It is often observed that the presence of a transgene copy of an endogenous gene silences the expression of both the endogenous gene and the transgene (EP 0 465 572 A1).

According to one embodiment of the invention accumulation of trehalose is brought about in plants wherein the capacity of producing trehalose has been introduced by introduction of a plant expressible gene construct encoding trehalose phosphate synthase (TPS).

Any trehalose phosphate synthase gene under the control of regulatory elements necessary for expression of DNA in plant cells, either specifically or constitutively, may be used, as long as it is capable of producing an active trehalose phosphate synthase activity. A preferred open reading frame according to the invention is one encoding a TPS-enzyme as represented in SEQIDNO: 2. It is well known that more than one DNA sequence may encode an identical enzyme, which fact is caused by the degeneracy of the genetic code. If desired, the open reading frame encoding the trehalose phosphate synthase activity may be adapted to codon usage in the host plant of choice, but this is not a requirement.

The isolated nucleic acid sequence represented by SEQIDNO: 2, may be used to identify trehalose phosphate synthase activities in other organisms and subsequently isolating and cloning them, by hybridizing DNA from other sources with a DNA- or RNA fragment obtainable from the *E. coli* gene. Preferably, such DNA sequences are screened by hybridizing under more or less stringent conditions (such as temperature and ionic strength of the hybridization mixture). Whether or not conditions are stringent also depends on the nature of the hybridization, *i.e.* DNA:DNA, DNA:RNA, RNA:RNA, as well as the length of the shortest hybridizing fragment. Those of skill in the art are readily capable of establishing a hybridization regime stringent enough to isolate TPS genes, while avoiding aspecific hybridization. As genes involved in trehalose synthesis from other sources become available these can be used in a similar way to obtain a plant expressible trehalose phosphate synthase gene according to the invention.

Sources for isolating trehalose phosphate synthase activities include microorganisms (*e.g.* bacteria, yeast, fungi), plants, animals, and the like. Isolated DNA sequences encoding trehalose phosphate activity from other sources may be used likewise in a method for producing trehalose according to the invention. As an example, genes for producing trehalose from yeast are disclosed in WO 93/17093.

The invention also encompasses nucleic acid sequences which have been obtained by modifying the nucleic acid sequence represented in SEQIDNO: 2 by mutating one or more codons so that it results in amino acid changes in the encoded protein, as long as mutation of the amino acid sequence does not entirely abolish trehalose phosphate synthase activity.

According to another embodiment of the invention, plants are

genetically altered to produce and accumulate trehalose in specific parts of the plant, which were selected on the basis of considerations such as substrate availability for the TPS-enzyme, insensitivity of the plant part to any putative adverse effects of trehalose on plant cell functioning, and the like. A preferred site of TPS- enzyme expression are starch storage parts of plants. In particular microtubers potato are considered to be suitable plant parts. A preferred promoter to achieve selective TPS-enzyme expression in microtubers and tubers of potato is obtainable from the region upstream of the open reading frame of the patatin gene of potato (*Solanum tuberosum*).

Plants may be further modified by introducing additional genes that encode phosphatases that are capable of the conversion of trehalose phosphate into trehalose. At least in potato tubers or micro-tubers, potato leaves and tobacco leaves and roots, endogenous phosphatase activity appears to be present, so that the introduction of a trehalose phosphate phosphatase (TPP) gene is not an absolute requirement.

According to another embodiment of the invention, trehalose accumulation is further enhanced by the inhibition of endogenous genes in order to enhance substrate availability for the trehalose phosphate synthase, as exemplified herein with the inhibition of endogenous sucrose phosphate synthase gene and the ADP-Glucose pyrophosphorylase gene (AGP-ase). Inhibition of undesired endogenous enzyme activity is achieved in a number of ways, the choice of which is not critical to the invention. Preferably gene inhibition is achieved through the so-called 'antisense approach'. Herein a DNA sequence is expressed which produces an RNA that is at least partially complementary to the RNA which encodes the enzymatic activity that is to be blocked (*e.g.* AGP-ase or SPS (Sucrose Phosphate Synthase), in the examples). It is preferred to use homologous antisense genes as these are more efficient than heterologous genes. The isolation of an antisense SPS gene from potato using a maize SPS-gene sequence as probe serves to illustrate the feasibility of this strategy. It is not meant to indicate that, for practicing the invention the use of homologous antisense fragments is required. An alternative method to block the synthesis of undesired enzymatic activities is the introduction into the genome of the plant host of an additional copy of an endogenous gene present in the plant host. It is often observed that such an additional copy of a gene silences

the endogenous gene: this effect is referred to in the literature as the co-suppressive effect, or co-suppression. Details of the procedure of enhancing substrate availability are provided in the Examples of WO 95/01446, incorporated by reference herein.

Preferred plant hosts among the *spermatophyta* are the *Angiospermae*, notably the *Dicotyledoneae*, comprising *inter alia* the *Solanaceae* as a representative family, and the *Monocotyledoneae*, comprising *inter alia* the *Gramineae* as a representative family. Suitable host plants, as defined in the context of the present invention include plants (as well as parts and cells of said plants) and their progeny which have been genetically modified using recombinant DNA techniques to cause or enhance production of trehalose in the desired plant or plant organ; these plants may be used directly (*e.g.* the plant species which produce edible parts) in processing or the trehalose may be extracted and/or purified from said host. Crops with edible parts according to the invention include those which have flowers such as cauliflower (*Brassica oleracea*), artichoke (*Cynara scolymus*), fruits such as apple (*Malus, e.g. domesticus*), banana (*Musa, e.g. acuminata*), berries (such as the currant, *Ribes, e.g. rubrum*), cherries (such as the sweet cherry, *Prunus, e.g. avium*), cucumber (*Cucumis, e.g. sativus*), grape (*Vitis, e.g. vinifera*), lemon (*Citrus limon*), melon (*Cucumis melo*), nuts (such as the walnut, *Juglans, e.g. regia;* peanut, *Arachis hypogeae*), orange (*Citrus, e.g. maxima*), peach (*Prunus, e.g. persica*), pear (*Pyra, e.g. communis*), pepper (*Solanum, e.g. capsicum*), plum (*Prunus, e.g. domestica*), strawberry (*Fragaria, e.g. moschata*), tomato (*Lycopersicon, e.g. esculentum*), leafs, such as alfalfa (*Medicago sativa*), cabbages (such as *Brassica oleracea*), endive (*Cichoreum, e.g. endivia*), leek (*Allium porrum*), lettuce (*Lactuca sativa*), spinach (*Spinaciaoleraceae*), tobacco (*Nicotiana tabacum*), roots, such as arrowroot (*Maranta arundinacea*), beet (*Beta vulgaris*), carrot (*Daucus carota*), cassava (*Manihot esculenta*), turnip (*Brassica rapa*), radish (*Raphanus sativus*), yam (*Dioscorea esculenta*), sweet potato (*Ipomoea batatas*) and seeds, such as bean (*Phaseolus vulgaris*), pea (*Pisum sativum*), soybean (*Glycin max*), wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), corn (*Zea mays*), rice (*Oryza sativa*), tubers, such as kohlrabi (*Brassica oleraceae*), potato (*Solanum tuberosum*), and the like. The edible parts may be conserved by drying in the presence of enhanced trehalose levels

produced therein due to the presence of a plant expressible trehalose phosphate synthase gene.

The method of introducing the plant expressible trehalose-phosphate synthase gene, or any other sense or antisense gene into a recipient plant cell is not crucial, as long as the gene is expressed in said plant cell. The use of *Agrobacterium tumefaciens or Agrobacterium rhizogenes* - mediated transformation is preferred, but other procedures are available for the introduction of DNA into plant cells. Examples are transformation of protoplasts using the calcium/polyethylene glycol method, electroporation microinjection and DNA-coated particle bombardment (Potrykus, 1990, Bio/Technol. 8, 535-542). Also combinations of *Agrobacterium* and coated particle bombardment may be used. Also transformation protocols involving other living vectors than *Agrobacterium* may be used, such as viral vectors (*e.g.* from the Cauliflower Mosaic Virus (CaMV) and or combinations of *Agrobacterium* and viral vectors, a procedure referred to as agroinfection (Grimsley N. *et al.*, 8 January 1987, Nature 325, 177-179). After selection and/or screening, the protoplasts, cells or plant parts that have been transformed are regenerated into whole plants, using methods known in the art (Horsch *et al.*, 1985, Science 225, 1229-1231).

The development of reproducible tissue culture systems for monocotyledonous crops, together with methods for introduction of genetic material into plant cells has facilitated transformation. Presently, preferred methods for transformation of monocot species are microprojectile bombardment of explants or suspension cells, and direct DNA uptake or electroporation (Shimamoto, *et al.*, 1989, Nature 338, 274-276). Transgenic maize plants have been obtained by introducing the *Streptomyces hygroscopicus bar*-gene, which encodes phosphinothricin acetyltransferase (an enzyme which inactivates the herbicide phosphinothricin), into embryogenic cells of a maize suspension culture by microprojectile bombardment (Gordon-Kamm, 1990, Plant Cell, 2, 603-618). The introduction of genetic material into aleurone protoplasts of other monocot crops such as wheat and barley has been reported (Lee, 1989, Plant Mol. Biol. 13, 21-30). Wheat plants have been regenerated from embryogenic suspension culture by selecting only the aged compact and nodular embryogenic callus tissues for the establishment of the embryogenic suspension cultures

(Vasil, 1990 Bio/Technol. 8, 429-434).

Monocotyledonous plants, including commercially important crops such as corn and rice may be obtained by *Agrobacterium*-mediated transformation according to Gould J, Michael D, Hasegawa O, Ulian EC, Peterson G, Smith RH, (1991) Plant. Physiol. 95, 426-434; Hiei Y. *et al.*, The Plant Journal 6(2), 271-282 and European patent 159 418 B1.

Suitable DNA sequences for control of expression of the plant expressible genes (including marker genes), such as transcriptional initiation regions, enhancers, non-transcribed leaders and the like, may be derived from any gene that is expressed in a plant cell. Also intended are hybrid promoters combining functional portions of various promoters, or synthetic equivalents thereof. Apart from constitutive promoters, inducible promoters, or promoters otherwise regulated in their expression pattern, *e.g.* developmentally or cell-type specific, may be used to control expression of the plant expressible genes according to the invention as long as they are expressed in plant parts that contain substrate for TPS.

To select or screen for transformed cells, it is preferred to include a marker gene linked to the plant expressible gene according to the invention to be transferred to a plant cell. The choice of a suitable marker gene in plant transformation is well within the scope of the average skilled worker; some examples of routinely used marker genes are the neomycin phosphotransferase genes conferring resistance to kanamycin (EP-B 131 623), the Glutathion-S-transferase gene from rat liver conferring resistance to glutathione derived herbicides (EP-A 256 223), glutamine synthetase conferring upon overexpression resistance to glutamine synthetase inhibitors such as phosphinothricin (WO87/05327), the acetyl transferase gene from *Streptomyces viridochromogenes* conferring resistance to the selective agent phosphinothricin (EP-A 275 957), the gene encoding a 5-enolshikimate-3- phosphate synthase (EPSPS) conferring tolerance to N-phosphonomethylglycine, the *bar* gene conferring resistance against Bialaphos (*e.g.* WO 91/02071) and the like. The actual choice of the marker is not crucial as long as it is functional (*i.e.* selective) in combination with the plant cells of choice.

The marker gene and the gene of interest do not have to be linked, since co-transformation of unlinked genes (U.S. Patent 4,399,216) is also an efficient process in plant transformation.

Preferred plant material for transformation, especially for dicotyledonous crops are leaf-discs which can be readily transformed and have good regenerative capability (Horsch R.B. *et al.*, (1985) Science 227, 1229-1231).

It is immaterial to the invention how the presence of two or more genes in the same plant is effected. This can *inter alia* done be achieved by one of the following methods:

(a) transformation of the plant line with a multigene construct containing more than one gene to be introduced,

(b) co-transforming different constructs to the same plant line simultaneously,

(c) subsequent rounds of transformation of the same plant with the genes to be introduced,

(d) crossing two plants each of which contains a different gene to be introduced into the same plant, or

(e) combinations thereof.

The field of application of the invention lies both in agriculture and horticulture, for instance due to improved properties of the modified plants as such (*e.g.* stress tolerance, such as cold tolerance, and preferably drought resistance, and increase in post-harvest quality and shelf-life of plants and plant products), as well as in any form of industry where trehalose is or will be applied in a process of forced water extraction, such as drying or freeze drying. Trehalose can be used or sold as such, for instance in purified form or in admixtures, or in the form of a plant product, such as tuber, a fruit, a flower containing the trehalose, either in native state or in (partially) dehydrated form, and the like. Plant parts harbouring (increased levels of) trehalose phosphate or trehalose may be used or sold as such or processed without the need to add trehalose.

Also trehalose can be extracted and/or purified from the plants or plant parts producing it and subsequently used in an industrial process. In the food industries trehalose can be employed by adding trehalose to foods before drying. Drying of foods is an important method of preservation. Trehalose seems especially useful to conserve food products through conventional air-drying, and to allow for fast reconstitution upon addition of water of a high quality product (Roser *et al.*, July 1991,

Trends in Food Science and Technology, pp. 166-169). The benefits include retention of natural flavors/fragrances, taste of fresh product, and nutritional value (proteins and vitamins). It has been shown that trehalose has the ability to stabilize proteins *e.g.* vaccines, enzymes and membranes, and to form a chemically inert, stable glass. The low water activity of such thoroughly dried food products prevents chemical reactions, that could cause spoilage.

Field crops like corn, cassava, potato, sugar beet and sugarcane have since long been used as a natural source for bulk carbohydrate production (starches and sucrose). The production of trehalose in such crops, facilitated by genetic engineering of the trehalose-biosynthetic pathway into these plant species, would allow the exploitation of such engineered crops for trehalose production.

Trehalose is also used in drying or storage of biological macromolecules, such as peptides, enzymes, polynucleotides and the like.

All references cited in this specification are indicative of the level of skill in the art to which the invention pertains. All publications, whether patents or otherwise, referred to previously or later in this specification are herein incorporated by reference as if each of them was individually incorporated by reference. In particular WO 95/01446, cited herein, describing the production of trehalose in higher plants by genetic manipulation is herein incorporated by reference.

The Examples given below illustrate the invention and are in no way intended to indicate the limits of the scope of the invention.

Experimental

DNA manipulations

All DNA procedures (DNA isolation from *E.coli*, restriction, ligation, transformation, etc.) are performed according to standard protocols (Sambrook *et al.* (1989) Molecular Cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, CSH, New York).

Strains

In all examples E.coli K-12 strain DH5α is used for cloning. The *Agrobacterium tumefaciens* strains used for plant transformation experiments

are EHA 105 and MOG 101 (Hood *et al.* 1993, Trans. Research 2, 208-218)

Construction of Agrobacterium strain MOG101

A binary vector system (Hoekema A., Hirsch, P.R., Hooykaas, P.J.J., and Schilperoort, R.A. (1983) Nature 303, 179) is used to transfer gene constructs into potato and tobacco plants. The helper plasmid conferring the *Agrobacterium tumefaciens* virulence functions is derived from the octopine Ti-plasmid pTiB6. MOG101 is an *Agrobacterium tumefaciens* strain carrying a non-oncogenic Ti-plasmid (Koekman *et al.* 1982, *supra*) from which the entire T-region is deleted and substituted by a bacterial Spectinomycin resistance marker from transposon Tn1831 (Hooykaas *et al.*, 1980 Plasmid 4, 64-75). The Ti-plasmid pTiB6 contains two adjacent T-regions, TL (T-left) and TR (T-right). To obtain a derivative lacking the TL- and TR-regions, we constructed intermediate vector pMOG579. Plasmid pMOG579 is a pBR322 derivative which contains 2 Ti-plasmid fragments homologous to the fragments located left and right outside the T-regions of pTiB6. The 2 fragments are separated in pMOG579 by a 2.5 kb BamHI - HindIII fragment from transposon Tn1831 (Hooykaas *et al.*, 1980 Plasmid 4, 64-75) carrying the spectinomycin resistance marker. The plasmid is introduced into *Agrobacterium tumefaciens* strain LBA1010 [C58-C9 (pTiB6) = a cured C58 strain in which pTiB6 is introduced (Koekman *et al.* (1982), *supra*), by triparental mating from *E.coli*, using HB101 8pRK2013 as a helper. Transconjugants are selected for resistance to Rifampicin (20 mg/l) and spectinomycin (250 mg/l). A double recombination between pMOG579 and pTiB6 resulted in loss of carbenicillin resistance (the pBR322 marker) and deletion of the entire T-region. Of 5000 spectinomycin resistant transconjugants replica plated onto carbenicillin (100 mg/l) 2 are found sensitive. Southern analysis (not shown) showed that a double crossing over event had deleted the entire T-region. The resulting strain is called MOG101. This strain and its construction is analogous to strain GV2260 (Deblaere *et al.* 1985, Nucl. Acid Res. 13, 4777-4788).

An alternative helper strain for MOG101 is *e.g.* LBA4404; this strain can also suitably be used for introduction of a binary plasmid, such as pMOG799 and subsequent plant transformation. Other suitable helper strains are readily available.

Isolation of a patatin promoter/construction of pMOG546

A patatin promoter fragment is isolated from chromosomal DNA of *Solanum tuberosum* cv. Bintje using the polymerase chain reaction. A set of oligonucleotides, complementary to the sequence of the upstream region of the λpat21 patatin gene (Bevan, M., Barker, R., Goldsbrough, A., Jarvis, M., Kavanagh, T. and Iturriaga, G. (1986) Nucleic Acids Res. 14: 5564-5566), is synthesized consisting of the following sequences:

5′ AAG CTT ATG TTG CCA TAT AGA GTA G 3′ PatB33.2 (SEQIDNO:3)

5′ GTA GTT GCC ATG GTG CAA ATG TTC 3′ PatATG.2 (SEQIDNO:4)

These primers are used to PCR amplify a DNA fragment of 1123bp, using chromosomal DNA isolated from potato cv. Bintje as a template. The amplified fragment shows a high degree of similarity to the λpat21 patatin sequence and is cloned using EcoRI linkers into a pUC18 vector resulting in plasmid pMOG546.

Construction of pMOG 799

pMOG 799 harbours the TPS gene from *E. coli* under control of the double enhanced 35S Cauliflower Mosaic promoter. The construction of this binary vector is described in detail in International patent application PCT/EP94/02167, incorporated herein by reference. A sample of an E. coli strain harbouring pMOG799 has been deposited under the Budapest Treaty at the Centraal Bureau voor Schimmelcultures, Oosterstraat 1, P.O. Box 273, 3740 AG Baarn, The Netherlands, on Monday 23 August, 1993: the Accession Number given by the International Depositary Institution is CBS 430.93.

Construction of pMOG845.

Plasmid pMOG546 containing the patatin promoter is digested with NcoI-KpnI, incubated with *E. coli* DNA polymerase I in the presence of dATP and dCTP thereby destroying the NcoI and KpnI site and subsequently relegated. From the resulting vector a 1.1kb EcoRI-SmaI fragment containing the patatin promoter is isolated and cloned into pMOG798 (described in detail in PCT/EP94/02167) linearized with SmaI-EcoRI consequently exchanging the 35S CaMV promoter for the patatin promoter. The resulting vector is linearized with HindIII and ligated with the following oligonucleotide duplex:

	(HindIII) PstI	KpnI HindIII	
5′	AGCT CTGCAG TGA GGTACC A	3′	TCV 11 (SEQIDNO:5)
3′	GACGTC ACT CCATGG TTCGA	5′	TCV 12 (SEQIDNO:6)

After checking the orientation of the introduced oligonucleotide duplex, the resulting vector is linearized with PstI-HindIII followed by the insertion of a 950bp PstI-HindIII fragment harbouring the potato proteinase inhibitor II terminator (PotPiII) (An, G., Mitra, A., Choi, H.K., Costa, M.A., An, K., Thornburg, R. W. and Ryan, C.A. (1989) The Plant Cell 1: 115-122). The PotPiII terminator is isolated by PCR amplification using chromosomal DNA isolated from potato cv. Desiree as a template and the following set of oligonucleotides:

5′	GTACCCTGCAGTGTGACCCTAGAC	3′	TCV 15 (SEQIDNO:7)
5′	TCGATTCATAGAAGCTTAGAT	3′	TCV 16 (SEQIDNO:8)

The TPS expression cassette is subsequently cloned as a EcoRI-HindIII fragment into the binary vector pMOG402 resulting in pMOG845 (fig. 4). A sample of *E.coli* Dhα strain, harbouring pMOG845 has been deposited at the Centraal Bureau voor Schimmelcultures, Oosterstraat 1, P.O. Box 273, 3740 AG Baarn, The Netherlands, on January 4, 1995; the Accession Number given by the International Depositary Institution is CBS 101.95.

Triparental matings

The binary vectors are mobilized in triparental matings with the *E. coli* strain HB101 containing plasmid pRK2013 (Ditta G., Stanfield, S., Corbin, D., and Helinski, D.R. *et al.* (1980) Proc. Natl. Acad. Sci. USA 77, 7347) into *Agrobacterium tumefaciens* strain MOG101 or EHA105 and used for transformation.

Transformation of tobacco (*Nicotiana tabacum* SR1)

Tobacco is transformed by cocultivation of plant tissue with *Agrobacterium tumefaciens* strain MOG101 containing the binary vector of interest as described. Transformation is carried out using cocultivation of tobacco

(*Nicotiana tabacum* SR1) leaf disks as described by Horsch *et al.* 1985, Science 227, 1229-1231. Transgenic plants are regenerated from shoots that grow on selection medium containing kanamycin, rooted and transferred to soil.

Transformation of potato

Potato (*Solanum tuberosum* cv. Kardal) is transformed with the *Agrobacterium* strain EHA 105 containing the binary vector of interest. The basic culture medium is MS30R3 medium consisting of MS salts (Murashige, T. and Skoog, F. (1962) Physiol. Plan. 14, 473), R3 vitamins (Ooms *et al.* (1987) Theor. Appl. Genet. 73, 744), 30 g/l sucrose, 0.5 g/l MES with final pH 5.8 (adjusted with KOH) solidified when necessary with 8 g/l Daichin agar. Tubers of *Solanum tuberosum* cv. Kardal are peeled and surface sterilized by burning them in 96% ethanol for 5 seconds. Extinguish the flames in sterile water and cut slices of approximately 2 mm thickness. Disks are cut with a bore from the vascular tissue and incubated for 20 minutes in MS30R3 medium containing 1-5 $x10^8$ bacteria/ml of *Agrobacterium* EHA 105 containing the binary vector. Wash the tuber discs with MS30R3 medium and transfer them to solidified postculture medium (PM). PM consists of M30R3 medium supplemented with 3.5 mg/l zeatin riboside and 0.03 mg/l indole acetic acid (IAA). After two days, discs were transferred to fresh PM medium with 200 mg/l cefotaxim and 100 mg/l vancomycin. Three days later, the tuber discs are transferred to shoot induction medium (SIM) which consists of PM medium with 250 mg/l carbenicillin and 100 mg/l kanamycin. After 4-8 weeks, shoots emerging from the discs are excised and placed on rooting medium (MS30R3-medium with 100 mg/l cefotaxim, 50 mg/l vancomycin and 50 mg/l kanamycin). The shoots are propagated axenically by meristem cuttings.

Induction of micro-tubers

Stem segments of *in vitro* potato plants harbouring an auxiliary meristem are transferred to micro-tuber inducing medium. Micro-tuber inducing medium contains 1 X MS-salts supplemented with R3 vitamins, 0.5 g/l MES (final pH= 5.8, adjusted with KOH) and solidified with 8 g/l Daishin agar, 60 g/l sucrose and 2.5 mg/l kinetin. After 3 to 5 weeks of growth in the dark at 24°C, micro-tubers are formed.

Trehalose assay

Trehalose was determined quantitatively by anion exchange chromatography with pulsed amperometric detection. Extracts were prepared by adding 1 ml boiling water to 1 g frozen material which was subsequently heated for 15' at 100°C. Samples (25 µl) were analyzed on a Dionex DX-300 liquid chromatograph equipped with a 4 x 250 mm Dionex 35391 carbopac PA-1 column and a 4 x 50 mm Dionex 43096 carbopac PA-1 precolumn. Elution was with 100 mM NaOH at 1 ml/min. Sugars were detected with a pulsed amperometric detector (Dionex, PAD-2). Commercially available trehalose (Sigma) was used as a standard.

Isolation of Validamycin A

Validamycin A is isolated from Solacol, a commercial agricultural formulation (Takeda Chem. Indust., Tokyo) as described by Kendall *et al.* (1990) Phytochemistry, Vol. 29, No. 8, pp. 2525-2528. The procedure involves ion exchange chromatography (QAE-Sephadex A-25 (Pharmacia), bed vol. 10 ml, equilibration buffer 0.2 mM Na-Pi pH 7) from a 3% agricultural formulation of Solacol. Loading 1 ml of Solacol on the column and eluting with water in 7 fractions, practically all Validamycin is recovered in fraction 4.

Based on a 100% recovery, using this procedure, the concentration of Validamycin A was adjusted to 110^{-3} M in MS-buffer, for use in trehalose accumulation tests.

Alternatively, Validamycin A and B may be purified directly from *Streptomyces hygroscopicus* var. *limoneus*, as described by Iwasa T. *et al.*, 1971, in The Journal of Antibiotics 24(2), 119-123, the content of which is incorporated herein by reference.

Construction of pMOG1027

pMOG1027 harbours the trehalase gene from *Solanum tuberosum* cv. Kardal in the reversed orientation under control of the double enhanced 35S Cauliflower Mosaic promoter. The construction of this vector is very similar to the construction of pMOG799 and can be performed by any person skilled in the art. After mobilization of this binary vector by triparental mating to *Agrobacterium*, this strain can be used to transform plant cells and to generate transgenic plants having reduced levels of trehalase

activity.

Construction of pMOG1028

pMOG1028 harbours the trehalase gene from *Solanum tuberosum* cv. Kardal in the reversed orientation under control of the tuber specific patatin promoter. The construction of this vector is very similar to the construction of pMOG845 and can be performed by any person skilled in the art. After mobilization of this binary vector by triparental mating to *Agrobacterium*, this strain can be used in potato transformation experiments to generate transgenic plants having reduced levels of trehalase activity in tuber-tissue.

EXAMPLE 1

Cloning of a full length E. coli otsA gene

In E.*coli* trehalose phosphate synthase (TPS) is encoded by the otsA gene located in the operon otsBA. The cloning and sequence determination of the otsA gene is described in detail in Example I of PCT/EP94/02167, herein incorporated by reference. To effectuate its expression in plant cells, the open reading frame has been linked to the transcriptional regulatory elements of the CaMV 35S RNA promoter, the translational enhancer of the ALMV leader, and the transcriptional terminator of the NOS-gene, as described in greater detail in Example I of PCT/EP94/02167.

A binary vector, pMOG799 (Fig. 1), containing the plant expressible otsA gene and a the kanamycin resistance gene as selectable marker between T-DNA borders, is used to transform potato and tobacco.

EXAMPLE 2

Trehalose production in tobacco plants transformed with pMOG799

Tobacco leaf discs are transformed with the binary vector pMOG799 using Agrobacterium tumefaciens. Transgenic shoots are selected on kanamycin. Transgenic plants are transferred to the greenhouse to flower and set seed after selfing (S1). Seeds of these transgenic plants are surface sterilised and germinated *in vitro* on medium with Kanamycin. Kanamycin resistant seedlings and wild-type tobacco plants are transferred to MS-medium supplemented with 10^{-3} M Validamycin A. As a control, transgenic seedlings

and wild-type plants are transferred to medium without Validamycin A. Analysis of leaves and roots of plants grown on Validamycin A shows elevated levels of trehalose compared to the control plants (Table 1). No trehalose was detected in wild-type tobacco plants.

Table 1

	with Validamycin A		without Validamycin A	
	leaf	roots	leaf	roots
pMOG799.1	0.0081	0.0044	-	0.003
pMOG799.13	0.0110	0.0080	-	-
pMOG799.31	0.0008	0.0088	-	-
Wild-type SR1	-	-	-	-

EXAMPLE 3

Trehalose production in potato micro-tubers transformed with pMOG845

Potato *Solanum tuberosum* cv. Kardal tuber discs are transformed with *Agrobacterium tumefaciens* EHA105 harbouring the binary vector pMOG845. Transgenic shoots are selected on kanamycin. Micro-tubers (m-tubers) are induced on stem segments of transgenic and wild-type plants cultured on m-tuber inducing medium supplemented with 10^{-3} M Validamycin A. As a control, m-tubers are induced on medium without Validamycin A. M-tubers induced on medium with Validamycin A showed elevated levels of trehalose in comparison with m-tubers grown on medium without Validamycin A (Table 2). No trehalose was detected in wild-type m-tubers.

Table 2.

	Trehalose (% fresh weight)	
	+Validamycin A	-Validamycin A
845-2	0.016	-
845-4	-	-
845-8	0.051	-
845-13	0.005	-
845-22	0.121	-
845-25	0.002	-
wT Kardal	-	-

EXAMPLE 4

Trehalose production in hydrocultures of tobacco plants transformed with pMOG799

Seeds (S1) of selfed tobacco plants transformed with the binary vector pMOG799 are surface sterilised and germinated *in vitro* on MS20MS medium containing 50 µg/ml Kanamycin. Kanamycin resistant seedlings are transferred to soil and grown in a growth chamber (temp. 23°C, 16 hours of light/day). After four weeks, seedlings were transferred to hydrocultures with ASEF clay beads with approximately 450 ml of medium. The medium contains 40 g/l Solacol dissolved in nano-water buffered with 0.5 g/l MES to adjust to pH 6.0 which is sieved through a filter to remove solid particles. Essential salts are supplemented by adding POKON™ (1.5 ml/l). The following antibiotics are added to prevent growth of micro-organisms: 500µg/ml Carbenicillin, 40µg/ml Nystatin and 100µg/ml Vancomycin. As a control, transgenic seedlings and wild-type plants are transferred to medium without Solacol. Analysis of leaves of plants grown on Solacol shows elevated levels of trehalose compared to the control plants (Table 3). No trehalose was detected in wild-type tobacco plants.

Table 3

	Solacol	Trehalose (%w/w)
pMOG 799.1-1	+	0.008
pMOG 799.1-2	+	0.004
pMOG 799.1-3	-	-
pMOG 799.1-4	-	-
pMOG 799.1-5	+	0.008
pMOG 799.1-6	-	-
pMOG 799.1-7	+	0.005
pMOG 799.1-8	-	-
pMOG 799.1-9	-	-
pMOG 799.1-10	+	0.007
Wild-type SR1-1	-	-
Wild-type SR1-2	+	-
Wild-type SR1-3	-	-
Wild-type SR1-4	+	-

Example 5

Cloning of a full length cDNA encoding trehalase from potato tuber tissue

Using the amino acid sequence of the conserved regions of known trehalase genes (*E.coli*, Yeast, Rabbit, *B. mori*)(figure 4), four degenerated primers were designed:

GG(C/T)GGI(C/A)G(C/A)TT(CGT/TAA)IGA(GT/AG)T(A/C)TA(TTAT/CGGC)TGGGAC Tase24 (SEQIDNO:11)

GTICCIGGIGGICGITT(TAA/CGT)IGA(GT/AG)T Tase25 (SEQIDNO:12)

GGIGG(T/C)TGI(GA/CT)ICGI(TG/CA)IAG(A/G)TA(A/G)TA Tase26 (SEQIDNO:13)

I(C/G)(C/A)G(TTI/GC)(AT/CCATCC)(A/G)AAICCITC Tase27 (SEQIDNO:14)

Combinations of these primers in PCR experiments with genomic DNA and cDNA from *S. tuberosum* cv. Kardal leaf and tuber material respectively as template, resulted in several fragments of the expected length. A number of 190 bp. fragments obtained with the primer combination Tase24 and Tase 26 were subcloned into a pGEM T vector and sequenced. Several of the clones analyzed showed homology with known trehalase sequences. To exclude the isolation of non-plant derived trehalase sequences, Southern blot analysis was performed with gDNA from potato cv. Kardal. A number of clones isolated did not cross-hybridize with Kardal genomic DNA and were discarded. Two isolated clones were identical, gTase15.4 derived from a genomic PCR experiment and cTase5.2 derived from a PCR on cDNA, both showing hybridization in Southern blot analysis. One single hybridizing band was detected (EcoRI 1.5 Kb, HindIII 3 Kb and BamHI larger than 12 Kb) suggesting the presence of only one copy of the isolated PCR fragment.

A cDNA library was constructed out of poly A^+ RNA from potato tubers (cv. Kardal) using a Stratagene cDNA synthesis kit and the vector Lambda ZAPII. Recombinant phages (500.000) were screened with the radiolabeled cTase5.2 PCR fragment resulting in the identification of 3 positive clones. After purification, two clones were characterised with restriction enzymes revealing inserts of 2.15 and 2.3 kb respectively. Their nucleotide sequence was 100% identical. The nucleic acid sequence

of on of these trehalase cDNA clones from *Solanum tuberosum* including its open reading frame is depicted in seq ID no:9, while the aminoacid sequence derived from this nucleic acid sequence is shown in seq. ID no:10. A plasmid harbouring an insert comprising the genetic information coding for trehalase has been deposited under no. CBS 804.95 with the Centraal Bureau voor Schimmelcultures, Oosterstraat 1, P.O. Box 273, 3740 AG Baarn, the Netherlands on December 8, 1995.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: MOGEN International n.v.
(B) STREET: Einsteinweg 97
(C) CITY: LEIDEN
(D) STATE: Zuid-Holland
(E) COUNTRY: The Netherlands
(F) POSTAL CODE (ZIP): NL-2333 CB
(G) TELEPHONE: (31).(71).5258282
(H) TELEFAX: (31).(71).5221471

(ii) TITLE OF INVENTION: Enhanced accumulation of trehalose in plants

(iii) NUMBER OF SEQUENCES: 14

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: EP 95.200.008.1
(B) FILING DATE: 04-JAN-1995

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: EP 95.202.415.6
(B) FILING DATE: 07-SEP-1995

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1446 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Escherichia coli
(B) STRAIN: CLONE: 7F11

(viii) POSITION IN GENOME:
(B) MAP POSITION: 41-42'

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 19..1446
(D) OTHER INFORMATION: /product= "trehalose phosphate synthase"
/gene= "otsA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
GAGAAAATAA CAGGAGTG ATG ACT ATG AGT CGT TTA GTC GTA GTA TCT AAC        51
                    Met Thr Met Ser Arg Leu Val Val Val Ser Asn
                      1               5                  10

CGG ATT GCA CCA CCA GAC GAG CAC GCC GCC AGT GCC GGT GGC CTT GCC        99
Arg Ile Ala Pro Pro Asp Glu His Ala Ala Ser Ala Gly Gly Leu Ala
            15                  20                  25

GTT GGC ATA CTG GGG GCA CTG AAA GCC GCA GGC GGA CTG TGG TTT GGC       147
Val Gly Ile Leu Gly Ala Leu Lys Ala Ala Gly Gly Leu Trp Phe Gly
        30                  35                  40

TGG AGT GGT GAA ACA GGG AAT GAG GAT CAG CCG CTA AAA AAG GTG AAA       195
Trp Ser Gly Glu Thr Gly Asn Glu Asp Gln Pro Leu Lys Lys Val Lys
    45                  50                  55

AAA GGT AAC ATT ACG TGG GCC TCT TTT AAC CTC AGC GAA CAG GAC CTT       243
Lys Gly Asn Ile Thr Trp Ala Ser Phe Asn Leu Ser Glu Gln Asp Leu
60                  65                  70                  75

GAC GAA TAC TAC AAC CAA TTC TCC AAT GCC GTT CTC TGG CCC GCT TTT       291
Asp Glu Tyr Tyr Asn Gln Phe Ser Asn Ala Val Leu Trp Pro Ala Phe
                80                  85                  90

CAT TAT CGG CTC GAT CTG GTG CAA TTT CAG CGT CCT GCC TGG GAC GGC       339
His Tyr Arg Leu Asp Leu Val Gln Phe Gln Arg Pro Ala Trp Asp Gly
            95                  100                 105

TAT CTA CGC GTA AAT GCG TTG CTG GCA GAT AAA TTA CTG CCG CTG TTG       387
Tyr Leu Arg Val Asn Ala Leu Leu Ala Asp Lys Leu Leu Pro Leu Leu
        110                 115                 120

CAA GAC GAT GAC ATT ATC TGG ATC CAC GAT TAT CAC CTG TTG CCA TTT       435
Gln Asp Asp Asp Ile Ile Trp Ile His Asp Tyr His Leu Leu Pro Phe
    125                 130                 135

GCG CAT GAA TTA CGC AAA CGG GGA GTG AAT AAT CGC ATT GGT TTC TTT       483
Ala His Glu Leu Arg Lys Arg Gly Val Asn Asn Arg Ile Gly Phe Phe
140                 145                 150                 155

CTG CAT ATT CCT TTC CCG ACA CCG GAA ATC TTC AAC GCG CTG CCG ACA       531
Leu His Ile Pro Phe Pro Thr Pro Glu Ile Phe Asn Ala Leu Pro Thr
                160                 165                 170
```

```
TAT GAC ACC TTG CTT GAA CAG CTT TGT GAT TAT GAT TTG CTG GGT TTC      579
Tyr Asp Thr Leu Leu Glu Gln Leu Cys Asp Tyr Asp Leu Leu Gly Phe
            175                 180                 185

CAG ACA GAA AAC GAT CGT CTG GCG TTC CTG GAT TGT CTT TCT AAC CTG      627
Gln Thr Glu Asn Asp Arg Leu Ala Phe Leu Asp Cys Leu Ser Asn Leu
        190                 195                 200

ACC CGC GTC ACG ACA CGT AGC GCA AAA AGC CAT ACA GCC TGG GGC AAA      675
Thr Arg Val Thr Thr Arg Ser Ala Lys Ser His Thr Ala Trp Gly Lys
    205                 210                 215

GCA TTT CGA ACA GAA GTC TAC CCG ATC GGC ATT GAA CCG AAA GAA ATA      723
Ala Phe Arg Thr Glu Val Tyr Pro Ile Gly Ile Glu Pro Lys Glu Ile
220                 225                 230                 235

GCC AAA CAG GCT GCC GGG CCA CTG CCG CCA AAA CTG GCG CAA CTT AAA      771
Ala Lys Gln Ala Ala Gly Pro Leu Pro Pro Lys Leu Ala Gln Leu Lys
                240                 245                 250

GCG GAA CTG AAA AAC GTA CAA AAT ATC TTT TCT GTC GAA CGG CTG GAT      819
Ala Glu Leu Lys Asn Val Gln Asn Ile Phe Ser Val Glu Arg Leu Asp
            255                 260                 265

TAT TCC AAA GGT TTG CCA GAG CGT TTT CTC GCC TAT GAA GCG TTG CTG      867
Tyr Ser Lys Gly Leu Pro Glu Arg Phe Leu Ala Tyr Glu Ala Leu Leu
        270                 275                 280

GAA AAA TAT CCG CAG CAT CAT GGT AAA ATT CGT TAT ACC CAG ATT GCA      915
Glu Lys Tyr Pro Gln His His Gly Lys Ile Arg Tyr Thr Gln Ile Ala
    285                 290                 295

CCA ACG TCG CGT GGT GAT GTG CAA GCC TAT CAG GAT ATT CGT CAT CAG      963
Pro Thr Ser Arg Gly Asp Val Gln Ala Tyr Gln Asp Ile Arg His Gln
300                 305                 310                 315

CTC GAA AAT GAA GCT GGA CGA ATT AAT GGT AAA TAC GGG CAA TTA GGC     1011
Leu Glu Asn Glu Ala Gly Arg Ile Asn Gly Lys Tyr Gly Gln Leu Gly
                320                 325                 330

TGG ACG CCG CTT TAT TAT TTG AAT CAG CAT TTT GAC CGT AAA TTA CTG     1059
Trp Thr Pro Leu Tyr Tyr Leu Asn Gln His Phe Asp Arg Lys Leu Leu
            335                 340                 345

ATG AAA ATA TTC CGC TAC TCT GAC GTG GGC TTA GTG ACG CCA CTG CGT     1107
Met Lys Ile Phe Arg Tyr Ser Asp Val Gly Leu Val Thr Pro Leu Arg
        350                 355                 360

GAC GGG ATG AAC CTG GTA GCA AAA GAG TAT GTT GCT GCT CAG GAC CCA     1155
Asp Gly Met Asn Leu Val Ala Lys Glu Tyr Val Ala Ala Gln Asp Pro
    365                 370                 375
```

```
GCC AAT CCG GGC GTT CTT GTT CTT TCG CAA TTT GCG GGA GCG GCA AAC      1203
Ala Asn Pro Gly Val Leu Val Leu Ser Gln Phe Ala Gly Ala Ala Asn
380                 385                 390                 395

GAG TTA ACG TCG GCG TTA ATT GTT AAC CCC TAC GAT CGT GAC GAA GTT      1251
Glu Leu Thr Ser Ala Leu Ile Val Asn Pro Tyr Asp Arg Asp Glu Val
                400                 405                 410

GCA GCT GCG CTG GAT CGT GCA TTG ACT ATG TCG CTG GCG GAA CGT ATT      1299
Ala Ala Ala Leu Asp Arg Ala Leu Thr Met Ser Leu Ala Glu Arg Ile
            415                 420                 425

TCC CGT CAT GCA GAA ATG CTG GAC GTT ATC GTG AAA AAC GAT ATT AAC      1347
Ser Arg His Ala Glu Met Leu Asp Val Ile Val Lys Asn Asp Ile Asn
        430                 435                 440

CAC TGG CAG GAG TGC TTC ATT AGC GAC CTA AAG CAG ATA GTT CCG CGA      1395
His Trp Gln Glu Cys Phe Ile Ser Asp Leu Lys Gln Ile Val Pro Arg
    445                 450                 455

AGC GCG GAA AGC CAG CAG CGC GAT AAA GTT GCT ACC TTT CCA AAG CTT      1443
Ser Ala Glu Ser Gln Gln Arg Asp Lys Val Ala Thr Phe Pro Lys Leu
460                 465                 470                 475

GCG                                                                  1446
Ala
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 476 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Thr Met Ser Arg Leu Val Val Val Ser Asn Arg Ile Ala Pro Pro
 1               5                  10                  15

Asp Glu His Ala Ala Ser Ala Gly Gly Leu Ala Val Gly Ile Leu Gly
            20                  25                  30

Ala Leu Lys Ala Ala Gly Gly Leu Trp Phe Gly Trp Ser Gly Glu Thr
        35                  40                  45

Gly Asn Glu Asp Gln Pro Leu Lys Lys Val Lys Lys Gly Asn Ile Thr
    50                  55                  60

Trp Ala Ser Phe Asn Leu Ser Glu Gln Asp Leu Asp Glu Tyr Tyr Asn
65                  70                  75                  80
```

Gln Phe Ser Asn Ala Val Leu Trp Pro Ala Phe His Tyr Arg Leu Asp
85 90 95

Leu Val Gln Phe Gln Arg Pro Ala Trp Asp Gly Tyr Leu Arg Val Asn
100 105 110

Ala Leu Leu Ala Asp Lys Leu Leu Pro Leu Leu Gln Asp Asp Asp Ile
115 120 125

Ile Trp Ile His Asp Tyr His Leu Leu Pro Phe Ala His Glu Leu Arg
130 135 140

Lys Arg Gly Val Asn Asn Arg Ile Gly Phe Phe Leu His Ile Pro Phe
145 150 155 160

Pro Thr Pro Glu Ile Phe Asn Ala Leu Pro Thr Tyr Asp Thr Leu Leu
165 170 175

Glu Gln Leu Cys Asp Tyr Asp Leu Leu Gly Phe Gln Thr Glu Asn Asp
180 185 190

Arg Leu Ala Phe Leu Asp Cys Leu Ser Asn Leu Thr Arg Val Thr Thr
195 200 205

Arg Ser Ala Lys Ser His Thr Ala Trp Gly Lys Ala Phe Arg Thr Glu
210 215 220

Val Tyr Pro Ile Gly Ile Glu Pro Lys Glu Ile Ala Lys Gln Ala Ala
225 230 235 240

Gly Pro Leu Pro Pro Lys Leu Ala Gln Leu Lys Ala Glu Leu Lys Asn
245 250 255

Val Gln Asn Ile Phe Ser Val Glu Arg Leu Asp Tyr Ser Lys Gly Leu
260 265 270

Pro Glu Arg Phe Leu Ala Tyr Glu Ala Leu Leu Glu Lys Tyr Pro Gln
275 280 285

His His Gly Lys Ile Arg Tyr Thr Gln Ile Ala Pro Thr Ser Arg Gly
290 295 300

Asp Val Gln Ala Tyr Gln Asp Ile Arg His Gln Leu Glu Asn Glu Ala
305 310 315 320

Gly Arg Ile Asn Gly Lys Tyr Gly Gln Leu Gly Trp Thr Pro Leu Tyr
325 330 335

Tyr Leu Asn Gln His Phe Asp Arg Lys Leu Leu Met Lys Ile Phe Arg
340 345 350

Tyr Ser Asp Val Gly Leu Val Thr Pro Leu Arg Asp Gly Met Asn Leu
355 360 365

Val Ala Lys Glu Tyr Val Ala Ala Gln Asp Pro Ala Asn Pro Gly Val
370 375 380

Leu Val Leu Ser Gln Phe Ala Gly Ala Ala Asn Glu Leu Thr Ser Ala
385 390 395 400

Leu Ile Val Asn Pro Tyr Asp Arg Asp Glu Val Ala Ala Ala Leu Asp
405 410 415

Arg Ala Leu Thr Met Ser Leu Ala Glu Arg Ile Ser Arg His Ala Glu
420 425 430

Met Leu Asp Val Ile Val Lys Asn Asp Ile Asn His Trp Gln Glu Cys
435 440 445

Phe Ile Ser Asp Leu Lys Gln Ile Val Pro Arg Ser Ala Glu Ser Gln
450 455 460

Gln Arg Asp Lys Val Ala Thr Phe Pro Lys Leu Ala
465 470 475

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAGCTTATGT TGCCATATAG AGTAG 25

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GTAGTTGCCA TGGTGCAAAT GTTC 24

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AGCTCTGCAG TGAGGTACCA 20

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GACGTCACTC CATGGTTCGA 20

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GTACCCTGCA GTGTGACCCT AGAC 24

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TCGATTCATA GAAGCTTAGA T 21

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2207 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Solanum tuberosum
(B) STRAIN: Kardal

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 161..1906

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 842..850
(D) OTHER INFORMATION: /function= "putative glycosylationsite"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CTTTTCTGAG TAATAACATA GGCATTGATT TTTTTTCAAT TAATAACACC TGCAAACATT 60

CCCATTGCCG GCATTCTCTG TTCTTACAAA AAAAAACATT TTTTTGTTCA CATAAATTAG 120

TTATGGCATC AGTATTGAAC CCTTTAACTT GTTATACAAT ATG GGT AAA GCT ATA 175
Met Gly Lys Ala Ile
1 5

```
ATT TTT ATG ATT TTT ACT ATG TCT ATG AAT ATG ATT AAA GCT GAA ACT      223
Ile Phe Met Ile Phe Thr Met Ser Met Asn Met Ile Lys Ala Glu Thr
                10                  15                  20

TGC AAA TCC ATT GAT AAG GGT CCT GTA ATC CCA ACA ACC CCT TTA GTG      271
Cys Lys Ser Ile Asp Lys Gly Pro Val Ile Pro Thr Thr Pro Leu Val
            25                  30                  35

ATT TTT CTT GAA AAA GTT CAA GAA GCT GCT CTT CAA ACT TAT GGC CAT      319
Ile Phe Leu Glu Lys Val Gln Glu Ala Ala Leu Gln Thr Tyr Gly His
        40                  45                  50

AAA GGG TTT GAT GCT AAA CTG TTT GTT GAT ATG TCA CTG AGA GAG AGT      367
Lys Gly Phe Asp Ala Lys Leu Phe Val Asp Met Ser Leu Arg Glu Ser
    55                  60                  65

CTT TCA GAA ACA GTT GAA GCT TTT AAT AAG CTT CCA AGA GTT GTG AAT      415
Leu Ser Glu Thr Val Glu Ala Phe Asn Lys Leu Pro Arg Val Val Asn
70                  75                  80                  85

GGT TCA ATA TCA AAA AGT GAT TTG GAT GGT TTT ATA GGT AGT TAC TTG      463
Gly Ser Ile Ser Lys Ser Asp Leu Asp Gly Phe Ile Gly Ser Tyr Leu
                90                  95                  100

AGT AGT CCT GAT AAG GAT TTG GTT TAT GTT GAG CCT ATG GAT TTT GTG      511
Ser Ser Pro Asp Lys Asp Leu Val Tyr Val Glu Pro Met Asp Phe Val
            105                 110                 115

GCT GAG CCT GAA GGC TTT TTG CCA AAG GTG AAG AAT TCT GAG GTG AGG      559
Ala Glu Pro Glu Gly Phe Leu Pro Lys Val Lys Asn Ser Glu Val Arg
        120                 125                 130

GCA TGG GCA TTG GAG GTG CAT TCA CTT TGG AAG AAT TTA AGT AGG AAA      607
Ala Trp Ala Leu Glu Val His Ser Leu Trp Lys Asn Leu Ser Arg Lys
    135                 140                 145

GTG GCT GAT CAT GTA TTG GAA AAA CCA GAG TTG TAT ACT TTG CTT CCA      655
Val Ala Asp His Val Leu Glu Lys Pro Glu Leu Tyr Thr Leu Leu Pro
150                 155                 160                 165

TTG AAA AAT CCA GTT ATT ATA CCG GGA TCG CGT TTT AAG GAG GTT TAT      703
Leu Lys Asn Pro Val Ile Ile Pro Gly Ser Arg Phe Lys Glu Val Tyr
                170                 175                 180

TAT TGG GAT TCT TAT TGG GTA ATA AGG GGT TTG TTA GCA AGC AAA ATG      751
Tyr Trp Asp Ser Tyr Trp Val Ile Arg Gly Leu Leu Ala Ser Lys Met
            185                 190                 195

TAT GAA ACT GCA AAA GGG ATT GTG ACT AAT CTG GTT TCT CTG ATA GAT      799
Tyr Glu Thr Ala Lys Gly Ile Val Thr Asn Leu Val Ser Leu Ile Asp
        200                 205                 210
```

```
CAA TTT GGT TAT GTT CTT AAC GGT GCA AGA GCA TAC TAC AGT AAC AGA      847
Gln Phe Gly Tyr Val Leu Asn Gly Ala Arg Ala Tyr Tyr Ser Asn Arg
    215                 220                 225

AGT CAG CCT CCT GTC CTG GCC ACG ATG ATT GTT GAC ATA TTC AAT CAG      895
Ser Gln Pro Pro Val Leu Ala Thr Met Ile Val Asp Ile Phe Asn Gln
230                 235                 240                 245

ACA GGT GAT TTA AAT TTG GTT AGA AGA TCC CTT CCT GCT TTG CTC AAG      943
Thr Gly Asp Leu Asn Leu Val Arg Arg Ser Leu Pro Ala Leu Leu Lys
                250                 255                 260

GAG AAT CAT TTT TGG AAT TCA GGA ATA CAT AAG GTG ACT ATT CAA GAT      991
Glu Asn His Phe Trp Asn Ser Gly Ile His Lys Val Thr Ile Gln Asp
            265                 270                 275

GCT CAG GGA TCA AAC CAC AGC TTG AGT CGG TAC TAT GCT ATG TGG AAT     1039
Ala Gln Gly Ser Asn His Ser Leu Ser Arg Tyr Tyr Ala Met Trp Asn
        280                 285                 290

AAG CCC CGT CCA GAA TCG TCA ACT ATA GAC AGT GAA ACA GCT TCC GTA     1087
Lys Pro Arg Pro Glu Ser Ser Thr Ile Asp Ser Glu Thr Ala Ser Val
    295                 300                 305

CTC CCA AAT ATA TGT GAA AAA AGA GAA TTA TAC CGT GAA CTG GCA TCA     1135
Leu Pro Asn Ile Cys Glu Lys Arg Glu Leu Tyr Arg Glu Leu Ala Ser
310                 315                 320                 325

GCT GCT GAA AGT GGA TGG GAT TTC AGT TCA AGA TGG ATG AGC AAC GGA     1183
Ala Ala Glu Ser Gly Trp Asp Phe Ser Ser Arg Trp Met Ser Asn Gly
                330                 335                 340

TCT GAT CTG ACA ACA ACT AGT ACA ACA TCA ATT CTA CCA GTT GAT TTG     1231
Ser Asp Leu Thr Thr Thr Ser Thr Thr Ser Ile Leu Pro Val Asp Leu
            345                 350                 355

AAT GCA TTC CTT CTG AAG ATG GAA CTT GAC ATT GCC TTT CTA GCA AAT     1279
Asn Ala Phe Leu Leu Lys Met Glu Leu Asp Ile Ala Phe Leu Ala Asn
        360                 365                 370

CTT GTT GGA GAA AGT AGC ACG GCT TCA CAT TTT ACA GAA GCT GCT CAA     1327
Leu Val Gly Glu Ser Ser Thr Ala Ser His Phe Thr Glu Ala Ala Gln
    375                 380                 385

AAT AGA CAG AAG GCT ATA AAC TGT ATC TTT TGG AAC GCA GAG ATG GGG     1375
Asn Arg Gln Lys Ala Ile Asn Cys Ile Phe Trp Asn Ala Glu Met Gly
390                 395                 400                 405

CAA TGG CTT GAT TAC TGG CTT ACC AAC AGC GAC ACA TCT GAG GAT ATT     1423
Gln Trp Leu Asp Tyr Trp Leu Thr Asn Ser Asp Thr Ser Glu Asp Ile
                410                 415                 420
```

```
TAT AAA TGG GAA GAT TTG CAC CAG AAC AAG AAG TCA TTT GCC TCT AAT      1471
Tyr Lys Trp Glu Asp Leu His Gln Asn Lys Lys Ser Phe Ala Ser Asn
            425                 430                 435

TTT GTT CCG CTG TGG ACT GAA ATT TCT TGT TCA GAT AAT AAT ATC ACA      1519
Phe Val Pro Leu Trp Thr Glu Ile Ser Cys Ser Asp Asn Asn Ile Thr
        440                 445                 450

ACT CAG AAA GTA GTT CAA AGT CTC ATG AGC TCG GGC TTG CTT CAG CCT      1567
Thr Gln Lys Val Val Gln Ser Leu Met Ser Ser Gly Leu Leu Gln Pro
    455                 460                 465

GCA GGG ATT GCA ATG ACC TTG TCT AAT ACT GGA CAG CAA TGG GAT TTT      1615
Ala Gly Ile Ala Met Thr Leu Ser Asn Thr Gly Gln Gln Trp Asp Phe
470                 475                 480                 485

CCG AAT GGT TGG CCC CCC CTT CAA CAC ATA ATC ATT GAA GGT CTC TTA      1663
Pro Asn Gly Trp Pro Pro Leu Gln His Ile Ile Ile Glu Gly Leu Leu
                490                 495                 500

AGG TCT GGA CTA GAA GAG GCA AGA ACC TTA GCA AAA GAC ATT GCT ATT      1711
Arg Ser Gly Leu Glu Glu Ala Arg Thr Leu Ala Lys Asp Ile Ala Ile
            505                 510                 515

CGC TGG TTA AGA ACT AAC TAT GTG ACT TAC AAG AAA ACC GGT GCT ATG      1759
Arg Trp Leu Arg Thr Asn Tyr Val Thr Tyr Lys Lys Thr Gly Ala Met
        520                 525                 530

TAT GAA AAA TAT GAT GTC ACA AAA TGT GGA GCA TAT GGA GGT GGT GGT      1807
Tyr Glu Lys Tyr Asp Val Thr Lys Cys Gly Ala Tyr Gly Gly Gly Gly
    535                 540                 545

GAA TAT ATG TCC CAA ACG GGT TTC GGA TGG TCA AAT GGC GTT GTA CTG      1855
Glu Tyr Met Ser Gln Thr Gly Phe Gly Trp Ser Asn Gly Val Val Leu
550                 555                 560                 565

GCA CTT CTA GAG GAA TTT GGA TGG CCT GAA GAT TTG AAG ATT GAT TGC      1903
Ala Leu Leu Glu Glu Phe Gly Trp Pro Glu Asp Leu Lys Ile Asp Cys
                570                 575                 580

TAATGAGCAA GTAGAAAAGC CAAATGAAAC ATCATTGAGT TTTATTTTCT TCTTTTGTTA    1963

AAATAAGCTG CAATGGTTTG CTGATAGTTT ATGTTTTGTA TTACTATTTC ATAAGGTTTT    2023

TGTACCATAT CAAGTGATAT TACCATGAAC TATGTCGTTC GGACTCTTCA AATCGGATTT    2083

TGCAAAAATA ATGCAGTTTT GGAGAATCCG ATAACATAGA CCATGTATGG ATCTAAATTG    2143

TAAACAGCTT ACTATATTAA GTAAAAGAAA GATGATTCCT CTGCTTTAAA AAAAAAAAAA    2203

AAAA                                                                 2207
```

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 581 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

```
Met Gly Lys Ala Ile Ile Phe Met Ile Phe Thr Met Ser Met Asn Met
 1               5                  10                  15

Ile Lys Ala Glu Thr Cys Lys Ser Ile Asp Lys Gly Pro Val Ile Pro
            20                  25                  30

Thr Thr Pro Leu Val Ile Phe Leu Glu Lys Val Gln Glu Ala Ala Leu
        35                  40                  45

Gln Thr Tyr Gly His Lys Gly Phe Asp Ala Lys Leu Phe Val Asp Met
    50                  55                  60

Ser Leu Arg Glu Ser Leu Ser Glu Thr Val Glu Ala Phe Asn Lys Leu
65                  70                  75                  80

Pro Arg Val Val Asn Gly Ser Ile Ser Lys Ser Asp Leu Asp Gly Phe
                85                  90                  95

Ile Gly Ser Tyr Leu Ser Ser Pro Asp Lys Asp Leu Val Tyr Val Glu
            100                 105                 110

Pro Met Asp Phe Val Ala Glu Pro Glu Gly Phe Leu Pro Lys Val Lys
        115                 120                 125

Asn Ser Glu Val Arg Ala Trp Ala Leu Glu Val His Ser Leu Trp Lys
    130                 135                 140

Asn Leu Ser Arg Lys Val Ala Asp His Val Leu Glu Lys Pro Glu Leu
145                 150                 155                 160

Tyr Thr Leu Leu Pro Leu Lys Asn Pro Val Ile Ile Pro Gly Ser Arg
                165                 170                 175

Phe Lys Glu Val Tyr Tyr Trp Asp Ser Tyr Trp Val Ile Arg Gly Leu
            180                 185                 190

Leu Ala Ser Lys Met Tyr Glu Thr Ala Lys Gly Ile Val Thr Asn Leu
        195                 200                 205

Val Ser Leu Ile Asp Gln Phe Gly Tyr Val Leu Asn Gly Ala Arg Ala
    210                 215                 220
```

Tyr Tyr Ser Asn Arg Ser Gln Pro Pro Val Leu Ala Thr Met Ile Val
225 230 235 240

Asp Ile Phe Asn Gln Thr Gly Asp Leu Asn Leu Val Arg Arg Ser Leu
245 250 255

Pro Ala Leu Leu Lys Glu Asn His Phe Trp Asn Ser Gly Ile His Lys
260 265 270

Val Thr Ile Gln Asp Ala Gln Gly Ser Asn His Ser Leu Ser Arg Tyr
275 280 285

Tyr Ala Met Trp Asn Lys Pro Arg Pro Glu Ser Ser Thr Ile Asp Ser
290 295 300

Glu Thr Ala Ser Val Leu Pro Asn Ile Cys Glu Lys Arg Glu Leu Tyr
305 310 315 320

Arg Glu Leu Ala Ser Ala Ala Glu Ser Gly Trp Asp Phe Ser Ser Arg
325 330 335

Trp Met Ser Asn Gly Ser Asp Leu Thr Thr Thr Ser Thr Thr Ser Ile
340 345 350

Leu Pro Val Asp Leu Asn Ala Phe Leu Leu Lys Met Glu Leu Asp Ile
355 360 365

Ala Phe Leu Ala Asn Leu Val Gly Glu Ser Ser Thr Ala Ser His Phe
370 375 380

Thr Glu Ala Ala Gln Asn Arg Gln Lys Ala Ile Asn Cys Ile Phe Trp
385 390 395 400

Asn Ala Glu Met Gly Gln Trp Leu Asp Tyr Trp Leu Thr Asn Ser Asp
405 410 415

Thr Ser Glu Asp Ile Tyr Lys Trp Glu Asp Leu His Gln Asn Lys Lys
420 425 430

Ser Phe Ala Ser Asn Phe Val Pro Leu Trp Thr Glu Ile Ser Cys Ser
435 440 445

Asp Asn Asn Ile Thr Thr Gln Lys Val Val Gln Ser Leu Met Ser Ser
450 455 460

Gly Leu Leu Gln Pro Ala Gly Ile Ala Met Thr Leu Ser Asn Thr Gly
465 470 475 480

Gln Gln Trp Asp Phe Pro Asn Gly Trp Pro Pro Leu Gln His Ile Ile
485 490 495

Ile Glu Gly Leu Leu Arg Ser Gly Leu Glu Glu Ala Arg Thr Leu Ala
500 505 510

Lys Asp Ile Ala Ile Arg Trp Leu Arg Thr Asn Tyr Val Thr Tyr Lys
515 520 525

Lys Thr Gly Ala Met Tyr Glu Lys Tyr Asp Val Thr Lys Cys Gly Ala
530 535 540

Tyr Gly Gly Gly Gly Glu Tyr Met Ser Gln Thr Gly Phe Gly Trp Ser
545 550 555 560

Asn Gly Val Val Leu Ala Leu Leu Glu Glu Phe Gly Trp Pro Glu Asp
565 570 575

Leu Lys Ile Asp Cys
580

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 6
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 15
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGYGGNMGMT TYRWNGARKT MTAYKRYTGG GAC 33

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 3
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 6
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 9
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 12
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 15
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 21
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GTNCCNGGNG GNCGNTTYRW NGARKT 26

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 3
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 9
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 12
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 15
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 18
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GGNGGYTGNS WNCGNYRNAG RTARTA 26

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 1
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 7
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 19
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 22
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

NSCRTTNRYC CATCCRAANC CNTC 24

CLAIMS

1. A process for producing trehalose in plant cells capable of producing trehalase by growing plant cells having the genetic information required for the production of trehalose and trehalase, or cultivating a plant or a part thereof comprising such plant cells, characterised in that said plant cells are grown, or said plant or a part thereof, is cultivated in the presence of a trehalase inhibitor.

2. A process according to claim 1, wherein said plant cells have been genetically altered so as to contain a chimeric trehalose phosphate synthase gene in a plant expressible form, preferably wherein the trehalose phosphate synthase gene comprises an open reading frame encoding trehalose phosphate synthase from *E. coli* in plant expressible form, more preferably wherein the open reading frame encoding trehalose phosphate synthase from *E. coli* is downstream of the CaMV 35S RNA promoter or the potato patatin promoter.

3. A process according to claim 1 or 2, wherein a *Solanum tuberosum* plant is cultivated, preferably wherein said plant has micro-tubers.

4. A process according to claim 3, wherein said plant is cultivated *in vitro*.

5. A process according to any one of claims 1 to 4, wherein said trehalase inhibitor comprises validamycin A in a form suitable for uptake by said plant cells, said plant, or a part thereof, preferably wherein the concentration of validamycin A is between 100 nM and 10 mM, more preferably between 0.1 and 1 mM, in aqueous solution.

6. A process according to any one of claims 1 to 4, wherein said trehalase inhibitor comprises the 86kD protein of the cockroach (*Periplaneta americana*) in a form suitable for uptake by said plant cells, said plant, or a part thereof.

7. A process according to any one of claims 1 to 4, wherein said plant cells have been genetically altered to contain the genetic information for a trehalase inhibitor, preferably wherein the trehalase inhibitor is the antisense gene to the gene encoding the information for trehalase or wherein the trehalase inhibitor is the 86kD protein of the American cockroach (*Periplaneta americana*).

8. A process according to any one of claims 1 to 7, wherein a plant, or a part thereof, accumulates trehalose in an amount above 0.01 % (fresh weight).

9. A plant, or a part thereof, or plant cells, obtainable by a process according to any one of the claims 1 to 8, which contain trehalose in an amount above 0.01% (fresh weight), preferably wherein said plant, or a part thereof is a *Solanaceae* species, more preferably *Solanum tuberosum* or *Nicotiana tabacum*.

10. A plant part according to claim 9, which is a tuber or a micro-tuber.

11. Tuber or micro-tubers of *Solanum tuberosum* containing trehalose.

12. Use of a plant, or plant part, according to claim 9 for extracting trehalose.

13. Use of a plant, or plant part, according to claim 9 in a process of forced extraction of water from said plant or plant part.

14. A plant according to claim 9, which has an increased stress tolerance, preferably increased drought tolerance.

15. A chimaeric plant expressible gene comprising in sequence a transcription initiation region obtainable from a gene, preferentially expressed in a plant part, particularly the patatin gene from *Solanum tuberosum*, a 5'-untranslated leader, an open reading frame encoding a trehalose phosphate synthase activity, and downstream of said open

reading frame a transciptional terminator region, preferably wherein said transcriptional terminator region is obtainable from the proteinase inhibitor-II gene of *Solanum tuberosum*.

16. A vector comprising a chimaeric plant expressible gene according to claim 15.

17. A recombinant plant genome comprising a chimaeric gene according to claim 16.

18. A plant cell having a recombinant genome according to claim 17.

19. A plant or a part thereof, consisting essentially of cells according to claim 18, wherein said plant is *Solanum tuberosum*.

20. A plant part according to claim 19, which is a tuber or a micro-tuber.

21. A process for obtaining trehalose, comprising the steps of growing plant cells according to claim 18, or cultivating a plant according to claim 19, or cultivating a plant part according to any one of claims 19 or 20, extracting trehalose from said plant cells, plants or parts.

22. A process for obtaining trehalose, comprising the steps of producing trehalose in plant cells, a plant or a part thereof, according to a process of any one of claims 1 to 8, and separating or extracting trehalose from said plant cells, plant or part thereof.

ABSTRACT

The invention provides a process for producing trehalose in plant cells capable of producing trehalase by growing plant cells having the genetic information required for the production of trehalose and trehalase, or cultivating a plant or a part thereof comprising such plant cells, characterised in that said plant cells are grown, or said plant or a part thereof, is cultivated in the presence of a trehalase inhibitor.